The Nature of Sampling Variability in the Index on Biotic Integrity (IBI) in Ohio Streams

Edward T. Rankin and Chris O. Yoder
Ecological Assessment Section
Division of Water Quality Planning and Assessment
Ohio EPA
1030 King Avenue
Columbus, OH 43212

Abstract
The Index of Biotic Integrity (IBI) was examined from a number of Ohio streams to determine the amount of variation that can be expected among replicate samples within years and among rivers with various degrees of cultural impact. Biosurvey data have been collected with standardized, pulsed-DC electrofishing techniques by the Ohio EPA over the past 12 years as part of its surface water monitoring program. The variation among samples, as measured by the coefficient of variation (CV) is lowest in streams and rivers that are least impacted by pollution (CV values < 10-12%) and increases in streams as cultural pollution increases (CV values up to 30-40%) until impacts are so toxic that there are only minimal fish communities (IBI scores 12-15). Indeed, high variability among samples in a year was a characteristic of impacted waterbodies. Variability among sampling passes also increased with decreasing habitat quality as measured by the Qualitative Habitat Evaluation Index (QHEI). Precision in the IBI compared favorably to precision in toxicological studies and analytical chemistry results. Among these approaches, as a direct measure of aquatic life, the IBI will be the most accurate arbiter of aquatic life use attainment in most situations.

Introduction

With increasing use of biosurvey data in state water resource monitoring programs it is important to understand, define and control the sources of variation common to biosurvey data. The Ohio EPA has been collecting fish community data, in a standardized manner, in streams and rivers since 1979 and has amassed data on over 3600 sites. This data provides an opportunity to examine patterns of data variability in response to temporal, geographical, and anthropogenic factors.

Five important sources of variability in biosurvey data are: (1) temporal variability (e.g., seasonal, daily, and diurnal changes in community composition), (2) sampling variability (e.g., related to gear, training, and effort), (3) spatial variability (e.g., related to stream size, faunal changes), (4) analytical variability (e.g., related to choice of the appropriate analytic tool), and (5) anthropogenic variability (e.g., degradation of water quality, habitat, toxic impacts to aquatic communities). It is critical to minimize or partition temporal, sampling, and analytical variation in biosurvey data to maximize the ability to distinguish anthropogenic impacts and variation. The goal of this paper is to define the "background" variation in the Index of Biotic Integrity (IBI) in minimally impacted streams (to define temporal and sampling variability) for comparison with variability in streams impacted by anthropogenic activities (i.e., those with aquatic life use impairment).

Background and Methods

The Ohio EPA uses pulsed-DC electrofishing methods (Ohio EPA 1989a) to capture a representative sample of the resident fish community in Ohio streams and rivers. Temporal variability in fish communities composition is minimized by sampling during daylight hours during the summer-early fall months (June 15 - October 15). In most situations we collect three sampling passes on different days during this period to detect within season (temporal) changes in the fish community. Recent work, however, in the largest Ohio rivers (Ohio River, lower Muskingum River) suggests that night sampling may provide more reliable results in these waterbodies (Sanders 1990) and we have excluded these rivers from this analysis.

Sampling variability is minimized through an extensive training program supported by a a detailed quality assurance manual (Ohio EPA 1989a) and the retention of experienced supervisory and field personnel (average experience > 10 years). Sampling equipment (longline, towboat, or boat mounted electrofishers) and methods and sampling effort are chosen to match the stream size and habitat (Figure 1). Effort is standardized on linear sampling distance which increases with stream size (Figure 1); minimum sampling times are defined for boat methods to ensure a minimum level of effort in large river habitats.

Macroinvertebrate community data and water column chemistry data are generally collected during the same time period as fish community data (June 15 - October 15). Field crews also perform habitat assessments with the Qualitative Habitat Evaluation Index (QHEI: Rankin 1989, Ohio EPA 1989a) within fish sampling zones. Water chemistry data, habitat data, knowledge of pollution sources, and biological response "signatures" (e.g., community response to different types of impacts) are used to determine the causes, sources, and magnitudes of impacts to aquatic life (Ohio EPA 1990). The Index of Biotic Integrity (IBI) is an analytical index used to assess fish community integrity; its applicability and derivation have been discussed elsewhere (Karr 1981, Fausch et al. 1984, Karr et al. 1986, Ohio EPA 1987a,b).

As a measure of variation we calculated the percent coefficient of variation (SD/Mean * 100) for the IBI at sites with three sampling passes

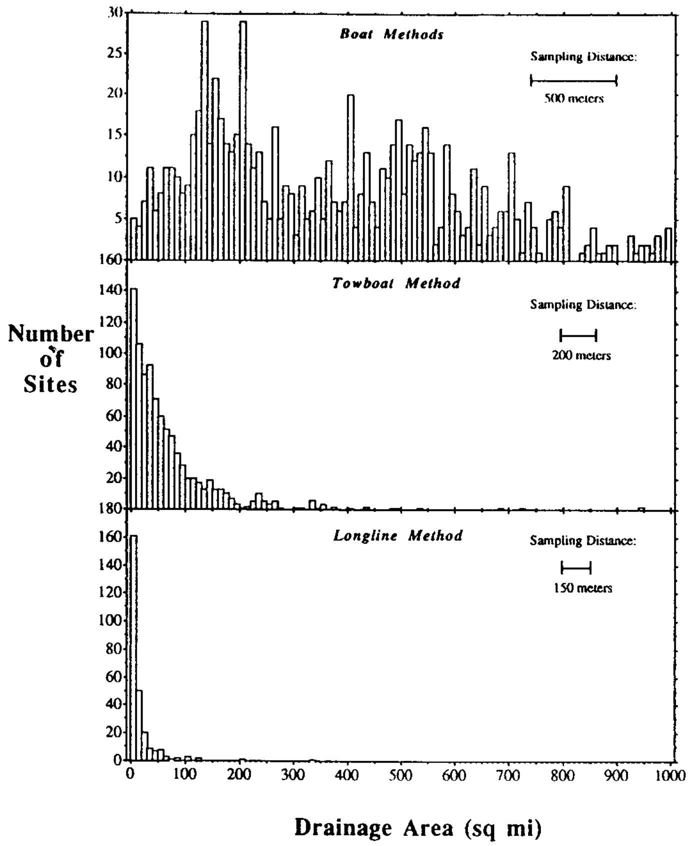

Figure 1. Range of stream sizes sampled by the Ohio EPA with boat, towboat, and longline pulsed-DC electrofishing methods. Sampling zone length for each method is included on each graph.

between June15 - October 15. The IBI, the Index of well-being (Iwb) for fish (Gammon et al. 1981) and the Invertebrate Community Index (ICI) for macroinvertebrates (Ohio EPA 1987b) comprise Ohio's biocriteria (Ohio Administrative Code 3745-1) and are the arbiter of aquatic life use impairment for Ohio's streams and rivers.

Although it is beyond the scope of this paper, one critical source of variation in water resource monitoring with biosurvey data is the appropriate choice of analytical tool. The advantages of broad-based, multimetric indices that have an ecological basis with both structural and functional components have been discussed by others (Karr 1981, Fausch et al. 1984, Karr et al. 1986, Karr 1990).

Results and Discussion

The median percent coefficient of variation (CV) at 1335 sites (1979-1989) with three sampling passes shows a distinct increase with decreasing IBI score except at the very lowest IBI range of 12-15 (Figure 2). Figure 2 is divided into IBI ranges that roughly correspond to Ohio's Exceptional Warmwater Habitat (EWH) aquatic life use IBI criteria, Warmwater Habitat (WWH) aquatic life use IBI criteria, and IBI scores that reflect impaired aquatic life uses. The median CV is generally less than 10% in EWH streams and 15% in WWH streams that achieve their respective IBI biocriteria. The distribution and range of CV values broadened significantly in streams with impaired aquatic life uses except at the very lowest IBI scores (12-15). By themselves increases in the variation of biosurvey data are an indication of impact to a stream. Cairns (1986) suggests that "...differences in variability rather than differences in averages or means might be the best measure of stress in natural systems".

Increases in variation are observed among streams affected by most types of impact (Figure 3). Ohio has no pristine, unimpacted streams. The "least impacted" streams in Ohio, however, such as the West Fork of Little Beaver Creek, Captina Creek, Rocky Fork of the Licking River, and the Kokosing River, have CV values of less than 5-10% and stable fish communities (as measured by the IBI). For example, the West Fork of Little Beaver Creek achieves an IBI of 50 or more (Ohio's EWH IBI criteria) in 25 of 27 sampling passes (Figure 4). This data spans five years and the two exceptions to this trend are due to a problem of recent origin.

Streams with impacted fish communities (IBI scores generally less than 40) had 75th percentile CV values of > 10-15% and as high as 30-40% (Figure 3). For example, the CV was negatively correlated with the QHEI (Qualitative Habitat Evaluation Index: Rankin 1989, Ohio EPA 1989), a measure of habitat quality (Figure 4). Low QHEI scores reflect low habitat quality that supports fewer habitat sensitive species and more tolerant individuals resulting in higher variability in catches and CPUE. Other impacts also resulted in increased variation in the IBI with toxic impacts among those associated with the highest IBI variation (Figure 3). Low species richness or low abundance of certain species, due to any impact type, increases the likelihood of IBI metrics being near scoring thresholds (1 vs 3 or 3 vs 5 points) and increases the variability in the IBI. Similarly, water quality impacts can reduce species numbers or affect trophic group composition through avoidance or mortality, and increase the variability of the IBI.

In contrast, extremely toxic impacts (IBI scores 12-15) were often characterized by little or no variation. In these situations few fish survive and metrics nearly always score a one. Exceptions are the downstream "edge" of a toxic effect (or episodic water quality impacts) which may shift the location of an impact over time, especially where there is migration from a nearby "refugia" with a healthy fish community. This situation was illustrated in Hurford Run near Canton Ohio (Figure 5). Upstream sections of Hurford Run had fish communities that were consistently very poor, but the fish community near the mouth fluctuated as tolerant, colonizing fish species (young-of-year green sunfish [Lepomis cyanellus], bluntnose minnow [Pimephales notatus], creek chub [Semotilus atromaculatus]) migrated from a mainstem "refugia".

The CV showed no regional pattern other than that which can be explained

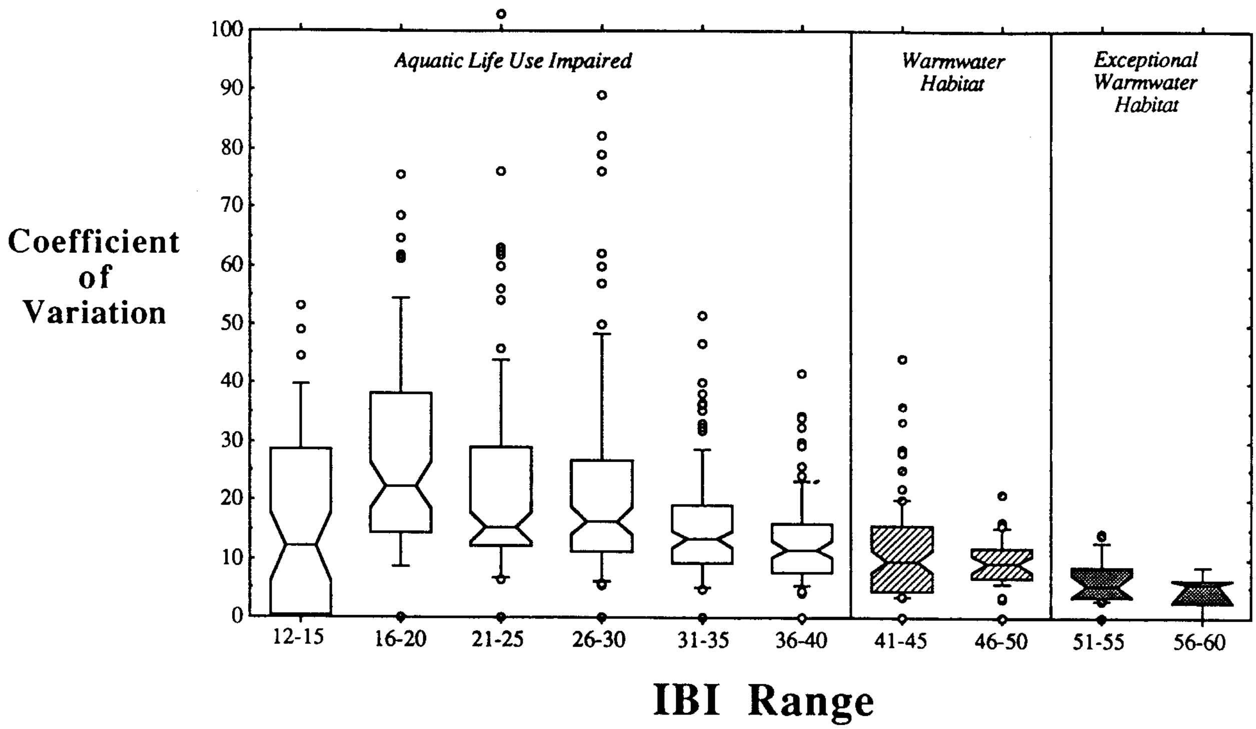

Figure 2. Median percent coefficient of variation (CV), 25th and 75th CV percentiles, CV range, and CV outliers (> 2 interquartile ranges from median) for the Index of Biotic Integrity (IBI) versus IBI range. CV values were calculated for sites with three sampling passes collected between June 15 - Oct 15. N = 1335 sites.

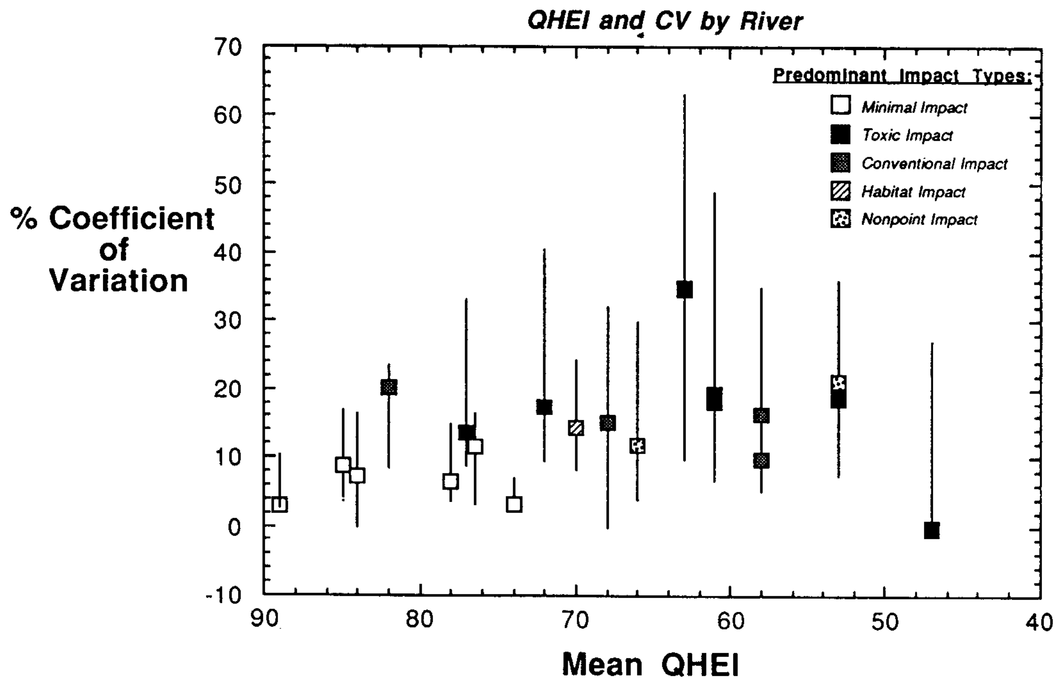

Figure 3. Median percent coefficient of variation (CV), and 10th and 90th CV percentiles for the Index of Biotic Integrity (IBI) versus QHEI (Qualitative Habitat Evaluation Index) for twenty Ohio streams and rivers. Shading of median values represents the predominant impact type in these streams.

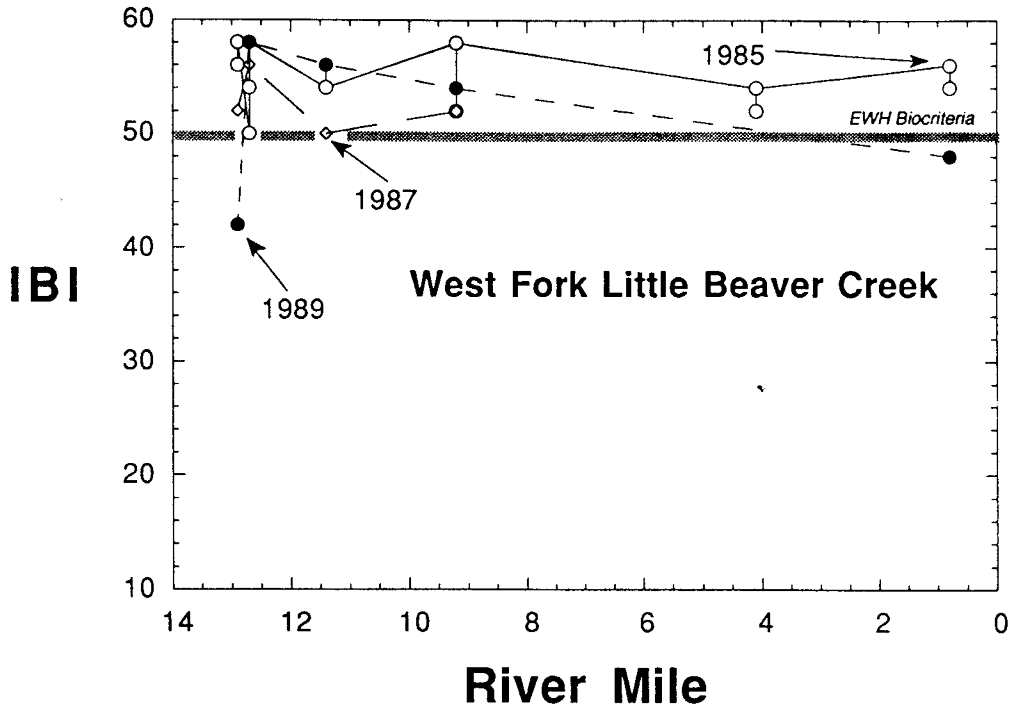

Figure 4. The Index of Biotic Integrity (IBI) versus river mile (upstream to downstream) for the West Fork of Little Beaver Creek (Columbiana Co., Ohio) for 1985 (N=3 passes), 1987 (N = 1 pass), and 1989 (N = 1 pass).

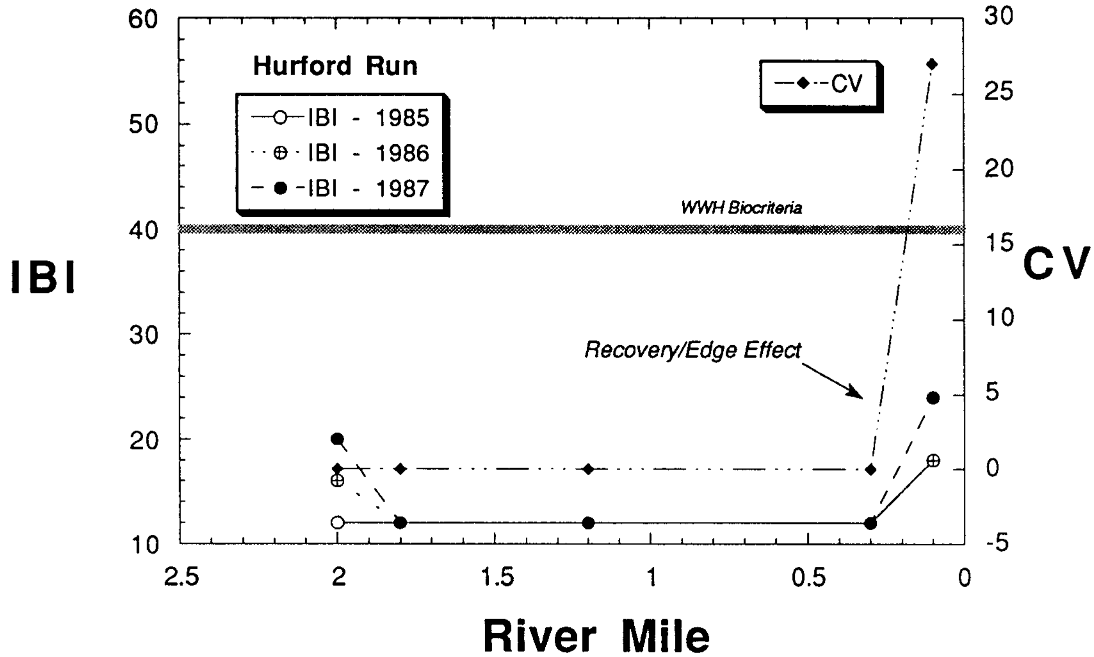

Figure 5. The Index of Biotic Integrity (IBI) and median percent coefficient of variation (CV) versus river mile (upstream to downstream) for the Hurford Run (Stark Co., Ohio) for 1985 (N=3 passes), 1986 (N = 1 pass), and 1988 (N = 1 pass).

by overall impacts within the ecoregions of Ohio (Figure 6). There was a slight trend in the upper threshold of variation in the IBI with stream size (Figure 7). Figure 7 represents the CV for streams in Ohio with IBI scores greater than 48 (i.e., the CV at these sites represents background variation due to inherent sampling variation and normal fluctuations in fish communities over time). The increase in the CV with stream size (Figure 7) most likely reflected the smaller proportion of the total community that was sampled in large versus small streams. Even in larger rivers, however, the CV was under 10-12% in the majority of situations.

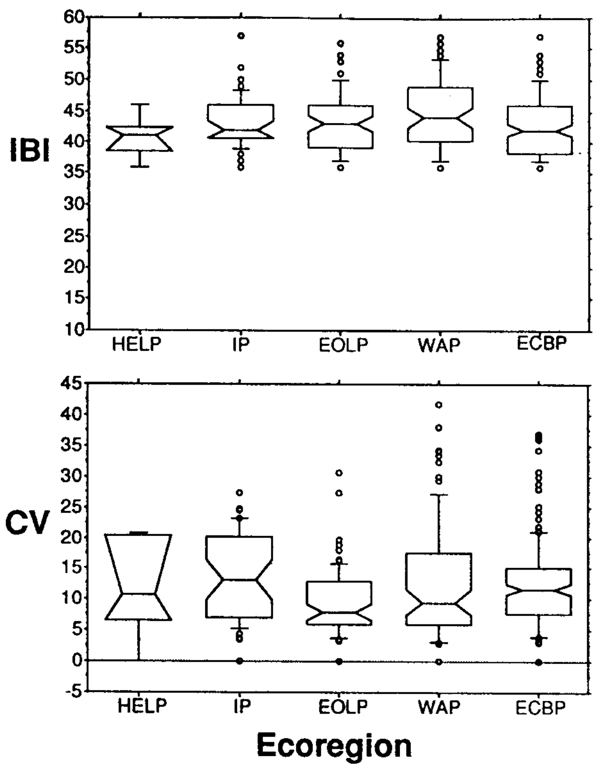

Figure 6. Boxplot of the median, 25th and 75th percentiles, range, and outliers (> 2 interquartile ranges from median) for the IBI (top panel) and percent coefficient of variation (CV, bottom panel) for sites in Ohio's five ecoregions. HELP: Huron Erie Lake Plain, IP: Interior Plateau, EOLP: Erie Ontario Lake Plain, WAP: Western Allegheny Plateau, ECBP: Eastern Corn Belt Plains.

Development of Ohio EPA "Significant Difference" in the IBI

Because we expected some background variation or "noise" in our samples we derived guidelines for detecting significant differences between IBI values from our intensive surveys and the regional reference sites used to derive our ecoregion-based biocriteria. We examined histograms of deviations in sample IBI values from mean IBI values at all locations where we had three sampling passes (Figure 8). We chose the 75th percentile value of this deviation from the mean as the limit of tolerable variation. This resulted in a guideline that the difference between a sample IBI and the ecoregion IBI biocriteria must be greater than 4 units to be classified as a significant departure. Because we used a mix of impacted and relatively unimpacted sites deviations of greater than 4 units probably reflects variation of anthropogenic origin. This is a protective criteria, however, because all available and applicable criteria for two organism groups (i.e., the modified iwb for fish in addition to the IBI and ICI for macroinvertebrates) must be met to fully attain an aquatic life use (Ohio EPA 1987b).

Detecting impacts and their underlying causes is more complex than simply determining significant departures from ecoregion biocriteria. For sites not attaining their aquatic life use the structural and functional characteristics of fish and macroinvertebrate communities provide information or "biological response signatures" about the type of impact that is affecting the aquatic life (Ohio EPA 1990a). Two sites that have similar IBI scores that indicate impaired communities may have very different community responses. The difference in the composition, function, and structure of the communities, in concert with chemical, toxicological, and physical data, provide clues to the cause or causes of impairment. Similarly, contrasts

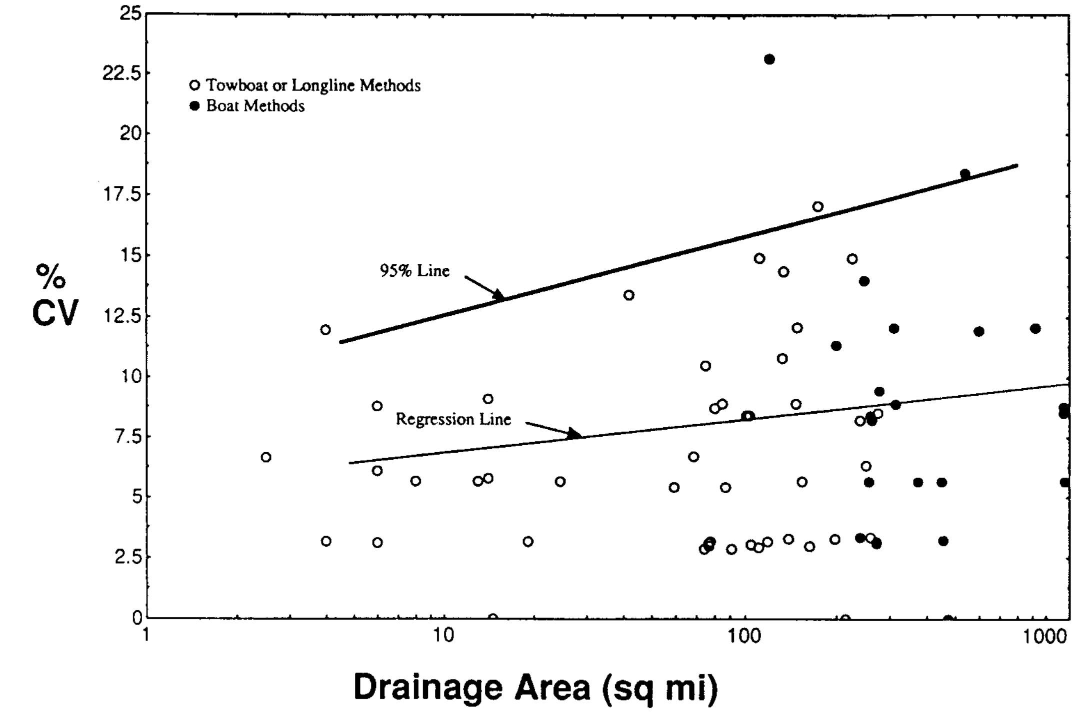

Figure 7. Median percent coefficient of variation (CV) versus drainage area for streams in Ohio with IBI scores > 48. The line on the graph represents an "upper threshold" and was drawn by eye through the upper 5% of the points.

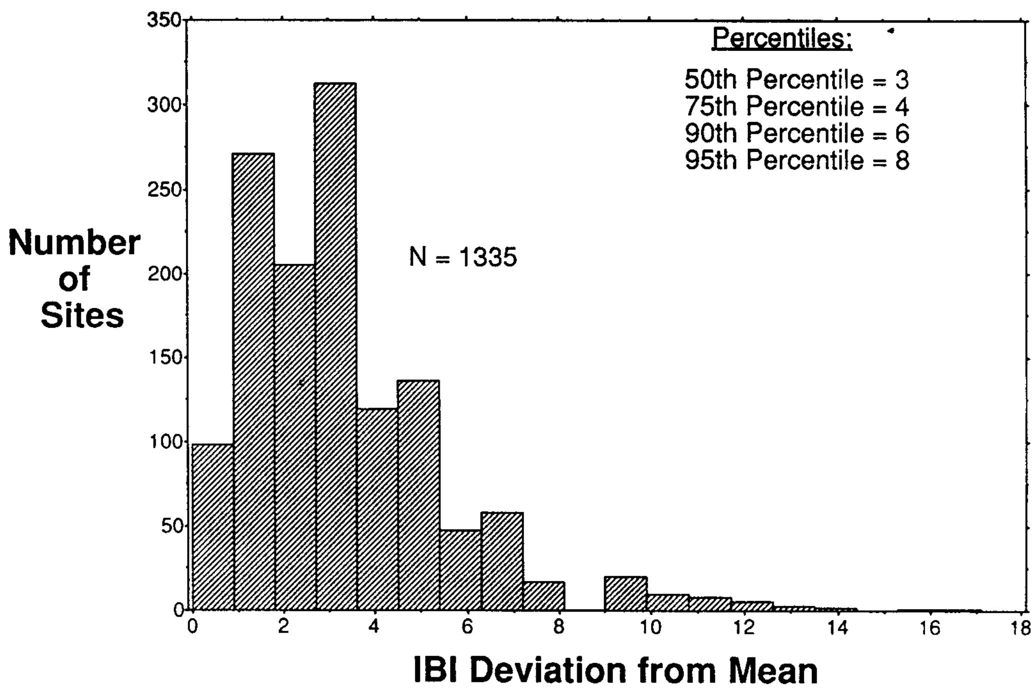

Figure 8. Frequency of the deviations of individual IBI passes from mean IBI values at stream sites with three sampling passes for all sites (solid bars) and reference sites (cross-hatched bars).

between the fish and macroinvertebrate community response are advantageous for detecting the type of impact. Work on formally classifying the responses of the biota to different types of impacts is in a developmental stage. New techniques, such as artificial intelligence (e.g., machine learning algorithms) may prove useful in this endeavor (David Davis, BBN Inc., personal communication).

Comparison of the CV values from biosurveys with other types of environmental monitoring data (e.g., water column chemistry, toxicity testing) provides additional perspective on the precision of the IBI. Mount (unpublished) compiled coefficient of variation values from a number of efforts to compare interlaboratory variability in toxicity testing and analytical water chemistry data. For organic and inorganic analyses most CV values were greater than 30% for the lower detection range of these parameters (e.g., mean of inorganic analyses = 125%). CV values, however, generally decrease when higher concentrations of compounds are analyzed (Turle 1990). The mean CV value (inter-laboratory variability) for toxicity tests (mostly LC50 values) was 30% (range: 0 - 66%; N = 16 CV values). Although replicate variability in the IBI was examined in this paper, the levels of interlaboratory variability associated with analytical chemistry data and toxicity testing are somewhat higher than the replicate biosurvey data. Though this interlaboratory variability is not strictly comparable to biosurvey replicate variability it does suggest that variability in biosurvey data is within or below the range of other, widely accepted environmental measurements.

CV values for replicate macroinvertebrate samples in a Wisconsin stream ranged from 6.2% to 43.6% (Szczykto 1989) depending on the index used in the analysis (all index scores were generated from the same data). Davis and Lubin (1989) calculated a CV of 20% for the Invertebrate Community Index (ICI) for all of the sites in Ohio EPA's regional reference site database. "Background" levels of precision are likely lower than 20% for replicate ICI scores for any given site because the reference sites are not homogeneous and represent a gradient of aquatic life performance. Nineteen replicate ICI scores at a relatively unimpacted test site in Big Darby Creek had a CV of 10.8%, which was lower than 8 of 9 of the index's underlying components. This CV value is similar to those found for the IBI in relatively unimpacted sites (see Figure 3).

Based on the data presented here the IBI scores collected by the Ohio EPA reflect low enough levels of sampling and natural variation to detect meaningful changes in biological integrity in streams. The precision of the IBI compares favorably with precision in analytical water chemistry methods and toxicity testing. However, this is not an effort to establish the "superiority" of one environmental measure over the other. Beyond considerations of precision, biosurvey data, water chemistry data and toxicity tests have specific applications where they are most appropriate and accurate. Our experience in Ohio has shown us that biosurvey, water chemistry, and toxicity testing are all necessary to completely and accurately define an impact to a stream in a complex situation, but that each is not necessarily independent of the other in all situations. There will be instances where one measure will carry more influence or weight than another. Unfortunately, this is not completely predictable at this point.

In the assessment of water resource impacts it is important to differentiate between accuracy and precision and to choose the appropriate "tool". Given an acceptable level of precision, emphasis should be put on environmental measures that accurately reflect water resource management goals (e.g., protection of aquatic life). For example, biological community data is free from assumptions and safety factors associated with laboratory derived data and accurately and directly reflects attainment of aquatic life uses (i.e., a high level of reality). Rankin and Yoder (1990) have shown that a reliance on water chemistry data and criteria alone under-estimated the impacts on aquatic life uses in Ohio in 49% of stream segments that were assessed. In contrast, only a small percentage of stream segments (< 3%) had biological communities that attained aquatic life uses, but violated chemical water quality criteria.

The IBI, when data collection methods are standardized, increases the accuracy of water resource assessments. Further work needs to: (1) identify biological response "signatures" for different types of impact, (2) identify situations where biosurvey data from multiple organism groups decreases the "variability" or increases the sensitivity of an assessment, (3) identify inter-laboratory variability in biosurvey data collection, and (4) compare variation between quantitative, standardized sampling methods (Ohio EPA approach described here) and more qualitative methods (e.g., Rapid Bioassessment Protocols, volunteer monitoring).

Acknowledgements
The work summarized here could not have been accomplished with the help of the staff biologists at the Ohio EPA: Marc Smith, Jeff DeShon, Chuck McKnight, Randy Sanders, Jack Freda, Roger Thoma, and Mike Bolton.

References

Fausch, K. D., J. R. Karr, and P. R. Yant. 1984. Regional application of an index of biotic integrity based on stream fish communities. Transactions of the American Fishery Society 113:39-55.

Karr, J. R., K. D. Fausch, P. L. Angermeier, P. R. Yant, and I. J. Schlosser. 1986. Assessing biological integrity in running waters: A method and its rationale. Illinois Natural History Survey Special Publication No. 5, 28 pp. Champaign, Illinois

Karr, J., R. 1989. Monitoring of biological integrity: An evolving approach to assessment and classification of water resources. Proceedings of the Midwest Pollution Control Biologists Meeting. U.S. EPA, Chicago, IL. EPA-905/9-89-007.

Mount, D. I. 1987 (unpublished). Comparison of test precision of effluent toxicity tests with chemical analyses. U. S. EPA, Environmental Research Laboratory, Duluth, Minnesota.

Ohio Environmental Protection Agency. 1987a. Biological criteria for the protection of aquatic life: Volume I. The role of biological data in water quality assessment. Division of Water Quality Planning and Assessment, Ecological Assessment Section, Columbus, Ohio.

Ohio Environmental Protection Agency. 1987b. Biological criteria for the protection of aquatic life: Volume II. Users manual for biological field assessment of Ohio surface waters. Division of Water Quality Planning and Assessment, Ecological Assessment Section, Columbus, Ohio.

Ohio Environmental Protection Agency. 1989a. Biological criteria for the protection of aquatic life: Volume III. Standardized biological field sampling and laboratory methods for assessing fish and macroinvertebrate communities. Division of Water Quality Planning and Assessment, Ecological Assessment Section, Columbus, Ohio.

Ohio Environmental Protection Agency. 1990a. Water Resource Inventory - Executive Summary: Volume I 1990 305(b) report. Edward T. Rankin, Chris Yoder, and Dennis Mishne. Division of Water Quality Planning and Assessment.

Rankin, E.T. and C.O. Yoder. 1990. A Comparison of Aquatic Life Use Impairment Detection and its Causes between an Integrated, Biosurvey-Based Environmental Assessment and its Water Column Chemistry Subcomponent. Appendix I In: Ohio Environmental Protection Agency. 1990. Water Resource Inventory - Executive Summary: Volume I 1990 305(b) report. Edward T. Rankin, Chris Yoder, and Dennis Mishne, editors. Division of Water Quality Planning and Assessment, Ecological Assessment Section, Columbus, Ohio.

Sanders, R. 1990. A 1989 night electrofishing survey of the Ohio River mainstem (RM 280.8 - 442.5). Ohio EPA, Division of Water Quality Planning and Assessment, Ecological Assessment Section, Columbus, Ohio

Szczytko, S. W. 1989. Variability of commonly used macroinvertebrate community metrics for assessing biomonitoring data and water quality in Wisconsin streams. Proceedings of the Midwest Pollution Control Biologists Meeting. U.S. EPA, Chicago, IL. EPA-905/9-89-007.

Turle, 1990. In-house reference materials as a means to QA: The CWS experience. Third Ecological Quality Assurance Workshop, Canada Centre for Inland Waters, Burlington, Ontario, Canada.